

Docket No. 1614.1069/HJS



IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Application of:

Megumi YOKOI et al.

Group Art Unit:

Serial No.:

Examiner:

Filed: August 25, 2000

For: MULTIPROCESSOR SYSTEM AND MEMORY ACCESS METHOD

**SUBMISSION OF CERTIFIED COPY OF PRIOR
FOREIGN APPLICATION IN ACCORDANCE WITH
THE REQUIREMENTS OF 37 C.F.R. § 1.55**

Assistant Commissioner for Patents
Washington, D.C. 20231

Sir:

In accordance with the provisions of 37 C.F.R. § 1.55, the applicant(s) submit(s)
herewith a certified copy of the following foreign application(s):

Japanese Patent Application No. 11-353729
Filed: December 13, 1999

It is respectfully requested that the applicant(s) be given the benefit of the foreign filing
date, as evidenced by the certified papers attached hereto, in accordance with the requirements
of 35 U.S.C. § 119.

Respectfully submitted,
STAAS & HALSEY LLP

Date: August 25, 2000

By: _____


H. J. Staas
Registration No. 22,010

700 Eleventh Street, N.W.
Suite 500
Washington, D.C. 20001
Telephone: (202) 434-1500
Facsimile: (202) 434-1501

日 本 国 特 許 庁
PATENT OFFICE
JAPANESE GOVERNMENT



別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出 願 年 月 日
Date of Application:

1999年12月13日

出 願 番 号
Application Number:

平成11年特許願第353729号

出 願 人
Applicant (s):

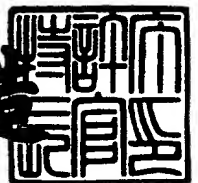
富士通株式会社

CERTIFIED COPY OF
PRIORITY DOCUMENT

2000年 7月28日

特許庁長官
Commissioner,
Patent Office

及川耕造



出証番号 出証特2000-3059579

【書類名】 特許願

【整理番号】 9951797

【提出日】 平成11年12月13日

【あて先】 特許庁長官 近藤 隆彦 殿

【国際特許分類】 G06F 12/00
G06F 13/00
G06F 15/16

【発明の名称】 メモリアクセス方法及びマルチプロセッサシステム

【請求項の数】 8

【発明者】

【住所又は居所】 神奈川県川崎市中原区上小田中4丁目1番1号 富士通株式会社内

【氏名】 横井 恵美

【発明者】

【住所又は居所】 神奈川県川崎市中原区上小田中4丁目1番1号 富士通株式会社内

【氏名】 和知 浩

【発明者】

【住所又は居所】 神奈川県川崎市中原区上小田中4丁目1番1号 富士通株式会社内

【氏名】 小田原 孝一

【発明者】

【住所又は居所】 神奈川県川崎市中原区上小田中4丁目1番1号 富士通株式会社内

【氏名】 渡部 徹

【発明者】

【住所又は居所】 神奈川県川崎市中原区上小田中4丁目1番1号 富士通株式会社内

【氏名】 村上 浩

【特許出願人】

【識別番号】 000005223

【氏名又は名称】 富士通株式会社

【代理人】

【識別番号】 100070150

【郵便番号】 150

【住所又は居所】 東京都渋谷区恵比寿4丁目20番3号 恵比寿ガーデン
プレイスタワー32階

【弁理士】

【氏名又は名称】 伊東 忠彦

【電話番号】 03-5424-2511

【手数料の表示】

【予納台帳番号】 002989

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【包括委任状番号】 9704678

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 メモリアクセス方法及びマルチプロセッサシステム

【特許請求の範囲】

【請求項 1】 データを保持するバッファと、各々がデータを一時的に保持するキャッシュメモリを有する複数のプロセッサを搭載したシステムモジュールが複数台、クロスバモジュールを介して接続された構成のマルチプロセッサシステムにおけるメモリアクセス方法であって、

任意の 1 台のシステムモジュール内のプロセッサからのリード要求に応答して、該任意の 1 台のシステムモジュール以外のシステムモジュールから先読みされたデータを、前記クロスバモジュール内のバッファに保持するステップを含む、メモリアクセス方法。

【請求項 2】 前記任意の 1 台のシステムモジュール内の 1 又は複数のプロセッサが実行するプログラムに応じて、任意のシステムモジュールに対するデータの先読みの可否を設定するステップを更に含む、請求項 1 記載のメモリアクセス方法。

【請求項 3】 前記先読みされたデータのデータ転送に、通常のデータ転送より優先度の低い優先度を付加するステップを更に含む、請求項 1 又は 2 記載のメモリアクセス方法。

【請求項 4】 複数台のシステムモジュールと、
少なくとも 1 台のクロスバモジュールと、
該システムモジュールと該クロスバモジュールとを接続するバスとを備え、
各システムモジュールは、データを保持するバッファと、各々がデータを一時的に保持するキャッシュメモリを有する複数のプロセッサと、システムモジュールに対するデータの入出力を制御する制御部とを有し、

該システムモジュール間のデータ転送は、該クロスバモジュールを介して行われ、

前記クロスバモジュールは、任意の 1 台のシステムモジュール内のプロセッサからのリード要求に応答して、該任意の 1 台のシステムモジュール以外のシステムモジュールから先読みされたデータを保持するバッファを有する、マルチプロ

セッサシステム。

【請求項 5】 前記任意の 1 台のシステムモジュールは、このシステムモジュール内の 1 又は複数のプロセッサが実行するプログラムに応じて、任意のシステムモジュールに対するデータの先読みの可否を設定する手段を更に有する、請求項 4 記載のマルチプロセッサシステム。

【請求項 6】 前記システムモジュールは、各々前記先読みされたデータのデータ転送に、通常のデータ転送より優先度の低い優先度を付加する手段を更に有する、請求項 4 又は 5 記載のマルチプロセッサシステム。

【請求項 7】 前記リード要求の要求アドレスのメモリを持つシステムモジュールは、このシステムモジュール内のキャッシュメモリの状態を知るより早いタイミングでデータの先読みを開始する手段を有する、請求項 4 ～ 6 のいずれか 1 項記載のマルチプロセッサシステム。

【請求項 8】 前記複数のシステムモジュールと、前記クロスバモジュールと、前記バスとはノードを構成し、

複数のノードが該クロスバモジュールを介して接続されている、請求項 4 ～ 7 のいずれか 1 項記載のマルチプロセッサシステム。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明はメモリアクセス方法及びマルチプロセッサシステムに係り、特に複数のプロセッサを搭載したシステムモジュールが複数、クロスバモジュールを介して接続された構成のマルチプロセッサシステムにおけるメモリアクセス方法、及び、このような構成のマルチプロセッサシステムに関する。

【0002】

【従来の技術】

従来のプロセッサシステムでは、1つのプロセッサからリード要求が出されると、このプロセッサのキャッシュメモリへのアクセスと同時に、メインメモリへデータの先読みアクセスを始める。キャッシュメモリへのアクセスがミスヒットすると、先読みアクセスによりメインメモリからバッファへ読み出しておいたデ

ータを使用することで、メモリアクセス時間の短縮を図るようにしている。

【0003】

従来のマルチプロセッサシステムでは、上記の如きプロセッサシステムが複数、バスを介して接続されている。従って、プロセッサのキャッシュメモリからデータを読み出す場合、バスを介して行われることが多い。

【0004】

【発明が解決しようとする課題】

マルチプロセッサシステムが大規模化するにつれて、データ転送の経路が非常に長くなり、従来のプロセッサシステムにおけるデータの先読みを単に適用したのでは、通常のデータ転送が妨害されたり、データの先読みによりバスが占有されて、マルチプロセッサシステム全体としての性能が低下してしまうという問題があった。

【0005】

又、先読みされたデータが保持されているバッファが、リード要求を行ったプロセッサから離れた場所にある場合には、先読みされたデータをリード要求を行ったプロセッサに転送するのに時間がかかり、先読みの本来のメリットが生かせないという問題もあった。

そこで、本発明は、リード要求を行ったプロセッサにできるだけ近い場所に先読みされたデータを保持して、通常のデータ転送を妨害することなく、先読みの本来のメリットを生かすことができ、マルチプロセッサシステム全体としての性能を向上可能なメモリアクセス方法及びマルチプロセッサシステムを提供することを目的とする。

【0006】

【課題を解決するための手段】

上記の課題は、請求項1記載の、データを保持するバッファと、各々がデータを一時的に保持するキャッシュメモリを有する複数のプロセッサを搭載したシステムモジュールが複数台、クロスバモジュールを介して接続された構成のマルチプロセッサシステムにおけるメモリアクセス方法であって、任意の1台のシステムモジュール内のプロセッサからのリード要求に応答して、該任意の1台のシス

テムモジュール以外のシステムモジュールから先読みされたデータを、前記クロスバモジュール内のバッファに保持するステップを含むメモリアクセス方法により達成できる。

【0007】

メモリアクセス方法は、請求項2記載の発明の如く、前記任意の1台のシステムモジュール内の1又は複数のプロセッサが実行するプログラムに応じて、任意のシステムモジュールに対するデータの先読みの可否を設定するステップを更に含んでも良い。つまり、データの先読みの可否は、プロセッサ毎に設定することも可能であり、1つのシステムモジュール内の全てでない複数のプロセッサに対して設定することも可能である。

【0008】

メモリアクセス方法は、請求項3記載の発明の如く、前記先読みされたデータのデータ転送に、通常のデータ転送より優先度の低い優先度を付加するステップを更に含んでも良い。

上記の課題は、請求項4記載の、複数台のシステムモジュールと、少なくとも1台のクロスバモジュールと、該システムモジュールと該クロスバモジュールとを接続するバスとを備え、各システムモジュールは、データを保持するバッファと、各々がデータを一時的に保持するキャッシュメモリを有する複数のプロセッサと、システムモジュールに対するデータの入出力を制御する制御部とを有し、該システムモジュール間のデータ転送は、該クロスバモジュールを介して行われ、前記クロスバモジュールは、任意の1台のシステムモジュール内のプロセッサからのリード要求に応答して、該任意の1台のシステムモジュール以外のシステムモジュールから先読みされたデータを保持するバッファを有するマルチプロセッサシステムによっても達成できる。

【0009】

請求項5記載の発明の如く、前記任意の1台のシステムモジュールは、このシステムモジュール内の1又は複数のプロセッサが実行するプログラムに応じて、任意のシステムモジュールに対するデータの先読みの可否を設定する手段を更に有する構成であっても良い。

請求項 6 記載の発明の如く、前記システムモジュールは、各々前記先読みされたデータのデータ転送に、通常のデータ転送より優先度の低い優先度を付加する手段を更に有する構成であっても良い。

【0 0 1 0】

請求項 7 記載の発明の如く、前記リード要求の要求アドレスのメモリを持つシステムモジュールは、このシステムモジュール内のキャッシュメモリの状態を知るより早いタイミングでデータの先読みを開始する手段を有する構成であっても良い。

請求項 8 記載の発明の如く、前記複数のシステムモジュールと、前記クロスバモジュールと、前記バスとはノードを構成し、マルチプロセッサシステムは、複数のノードが該クロスバモジュールを介して接続されている構成であっても良い。

【0 0 1 1】

従って、本発明によれば、リード要求を行ったプロセッサにできるだけ近い場所に先読みされたデータを保持して、通常のデータ転送を妨害することなく、先読みの本来のメリットを生かすことができ、マルチプロセッサシステム全体としての性能を向上可能となる。

【0 0 1 2】

【発明の実施の形態】

以下、本発明になるメモリアクセス方法及び本発明になるマルチプロセッサシステムの各実施例を、図面と共に説明する。

【0 0 1 3】

【実施例】

図 1 は、本発明になるマルチプロセッサシステムの第 1 実施例を示すブロック図である。マルチプロセッサシステムの第 1 実施例では、本発明になるメモリアクセス方法の第 1 実施例を採用する。

図 1 において、マルチプロセッサシステムは、大略複数のシステムモジュール（以下、システムボード（SB）と言う）1-1～1-Nと、クロスバモジュール（以下、クロスバボード（XB）と言う）2と、これらを接続するバス 3 とか

らなる。SB 1-1~1-Nは、夫々同じ構成を有し、Nは2以上の整数である。又、説明の便宜上、各SB 1-1~1-Nは2つのプロセッサを有するものとするが、3つ以上のプロセッサを有しても良い。

【0014】

各SB 1-1~1-Nは、キャッシュメモリ 11aを含むCPU等からなるプロセッサ 11、キャッシュメモリ 12aを含むプロセッサ 12、キャッシュメモリ 11のタグ 13、キャッシュメモリ 12のタグ 14、全体制御回路 15、メインメモリ 16、メモリアクセス制御装置 17、データ入出力制御装置 18、キャッシュ情報制御装置 19、バス調停装置 20及びデータ格納バッファ 21を有する。

【0015】

又、XB 2は、データ転送装置 31、アドレス広報装置 32及びキャッシュ情報広報装置 33を有する。データ転送装置 31は、データ格納バッファ 34とバス調停装置 35を有する。XB 2は、各SB 1-1~1-Nから送られてくる情報を選択、若しくは、マージして、バス 3を介して各SB 1-1~1-Nに送り返す機能を有する。

【0016】

通常のリード処理の手順は、次の通りである。即ち、①あるSBから発行されるリード要求は、XB 2によりバス 3を介して各SB 1-1~1-Nに供給され、②全てのSB 1-1~1-Nがキャッシュメモリの状態を示すキャッシュ情報をXB 2へ供給し、③XB 2が選択、若しくは、マージされたキャッシュ情報を各SB 1-1~1-Nに供給し、④どのSB 1-1~1-Nのキャッシュメモリにも有効なデータが無い、即ち、メインメモリ 16にある情報が最新の時は、対象アドレスのメモリを持つSBがメインメモリ 16へのアクセスを開始し、データを転送する。

【0017】

尚、各SB 1-1~1-Nにおいて、プロセッサ 11、12、メインメモリ 16等を除く回路部分は、1又は複数の大規模集積回路(LSI)により構成しても良い。

次に、本実施例の動作を説明する。SB 1-1 内のプロセッサ 11 においてミスヒットが発生すると、プロセッサ 11 はリード要求を SB 1-1 内の全体制御回路 15 に発行する。これに応答して、全体制御回路 15 は、バス 3 を介してリード要求を全ての SB 1-1 ~ 1-N に対して転送する。

【0018】

リード要求が転送された各 SB 1-1 ~ 1-N 内では、リード要求がキャッシュ情報制御装置 19 を介して全体制御回路 15 に供給される。つまり、リード要求は、要求元を含む全ての SB 1-1 ~ 1-N に転送される。キャッシュ情報制御装置 19 は、各々のキャッシュメモリ 11a, 12a のタグ 13, 14 を調べて、タグ情報をバス 3 を介して XB 2 内のキャッシュ情報広報装置 33 に出力する。又、全体制御回路 15 は、リード要求により要求されているアドレス（要求アドレス）が、全体制御回路 15 が属する SB 内のアドレスであるか否かを判定し、この SB 内のアドレスであればメモリアクセス制御装置 17 がメインメモリ 16 のアクセスを起動する。

【0019】

説明の便宜上、図 1 は、SB 1-2 が、SB 1-1 内のプロセッサ 11 のリード要求により要求されているアドレスのメモリを持っている場合を示すものとする。従って、リード要求を受け取った SB 1-2 内では、全体制御回路 15 が、要求アドレスのメモリが SB 1-2 内のメインメモリ 16 であることを判定すると共に、キャッシュ情報制御装置 19 は、各々のキャッシュメモリ 11a, 12a のタグ 13, 14 を調べて、タグ情報をバス 3 を介して XB 2 内のキャッシュ情報広報装置 33 に出力する。この時、全体制御回路 15 は、各々のキャッシュメモリ 11a, 12a のタグ情報をキャッシュ情報制御装置 19 から知らされる前に、メモリアクセス制御装置 17 を起動し、メモリアクセス制御装置 17 は先読みによるメインメモリ 16 のアクセスの起動を試みる。

【0020】

又、SB 1-2 内では、メモリアクセス制御装置 17 がデータ入出力制御装置 18 を起動し、データ入出力制御装置 18 はデータ保持バッファ 21 の準備をする。これにより、リード要求の要求アドレスのデータが、メインメモリ 16 から

データ保持バッファ 2 1 に転送される。尚、メインメモリ 1 6 が他からのアクセスにより起動できない場合には、メモリアクセス制御装置 1 7 は、起動が可能となるまでメインメモリ 1 6 へのアクセスを続ける。

【 0 0 2 1 】

更に、SB 1 - 2 内の全体制御回路 1 5 は、各々のキャッシュメモリ 1 1 a, 1 2 a のタグ情報をキャッシュ情報制御装置 1 9 から知らされると、メモリアクセス制御装置 1 7 の先読みアクセスを中断し、タグ情報を調査、即ち、解析する。全体制御回路 1 5 は、SB 1 - 2 内の全てのキャッシュメモリ 1 1 a, 1 2 a に要求アドレスがない場合にのみ、メモリアクセス制御装置 1 7 を再び起動し、メモリアクセス制御装置 1 7 はメインメモリ 1 6 へのアクセスを再開する。

【 0 0 2 2 】

SB 1 - 2 内で、メインメモリ 1 6 からデータ保持バッファ 2 1 に転送されたデータは、以下のようにして、最終的にはリード要求を発行したプロセッサ 1 1 が属する SB 1 - 1 内のデータ保持バッファ 2 1 に転送される。

具体的には、SB 1 - 2 内において、バス調停装置 2 0 は、通常 of データ転送のためのバス 3 の使用を妨げないように、データ保持バッファ 2 1 内に保持された、先読みアクセスにより読み出されたデータの優先度を通常 of データ転送の優先度より下げて、バス 3 の調停（アービトレーション）を行う。バス調停装置 2 0 が行うバス 3 の調停では、バス調停装置 2 0 が属する SB 内のどのプロセッサからの要求を出力するかを決定する。バス調停装置 2 0 がバス 3 の使用権を獲得すると、先読みデータを XB 2 のデータ転送装置 3 1 内のデータ保持バッファ 3 4 へ出力する。先読みデータは、XB 2 のデータ保持バッファ 3 4 に一旦保持され、バス調停装置 3 5 は、通常 of データ転送のためのバス 3 の使用を妨げないように、データ保持バッファ 3 4 内に保持された先読みデータの優先度を通常 of データ転送の優先度より下げて、バス 3 の調停を行う。バス調停装置 3 5 が行うバス 3 の調停では、各 SB 1 - 1 ~ 1 - N から XB 2 に供給されるどの要求を選択するかを決定する。バス調停装置 3 5 がバス 3 の使用権を獲得すると、先読みデータを、リード要求を発行したプロセッサ 1 1 が属する SB 1 - 1 内のデータ保持バッファ 2 1 へ出力する。これにより、先読みデータは、SB 1 - 1 内のデー

タ保持バッファ 2 1 に一旦保持される。

【0 0 2 3】

先読みデータには、例えばデータが先読みデータであることを示す先読みデータフラグを付加しておくことで、バス調停装置 2 0, 3 5 においてデータ転送の優先度を下げるか否かを判断できる。

各 S B 1 - 1 ~ 1 - N から出力されるキャッシュメモリ 1 1 a, 1 2 a に関するキャッシュ情報は、X B 2 内のキャッシュ情報広報装置 3 3 によりひとまとめにされ、各 S B 1 - 1 ~ 1 - N に全キャッシュ情報が転送される。全キャッシュ情報に基づいて、先読みが失敗したと判断された時点で、各データ保持バッファ 2 1, 3 4 内の先読みデータは破棄される。又、先読みが成功したと判断された時点で、先読みデータは正式なデータとみなされ、優先度は通常のデータ転送と同じ優先度に上げられて、処理が続けられる。

【0 0 2 4】

全キャッシュ情報の判断と判断結果の通知は、各 S B 1 - 1 ~ 1 - N 及び X B 2 毎に行っても、1 つの S B 又は X B 2 がまとめて行っても良い。X B 2 が全キャッシュ情報の判断と判断結果の通知を行う場合には、キャッシュ情報広報装置 3 3 が X B 2 内で全キャッシュ情報を広報するか、或いは、各 S B 1 - 1 ~ 1 - N 側から X B 2 内のデータ保持バッファ 3 4 に制御信号を出力するようにすることもできる。

【0 0 2 5】

これにより、より早いタイミングでの先読みの開始と、通常のデータ転送のためのバス 3 の使用を妨げない先読みデータの転送、及び先読み成功時のプロセッサへの先読みデータの素早い取り込みが可能となり、マルチプロセッサシステムにおいて先読みの本来のメリットを十分生かすことが可能となる。

次に、データの先読みを行うか否かの設定について、図 2 と共に説明する。本実施例では、データの先読みを行うか否かの設定は、例えば各プロセッサ 1 1, 1 2 内のレジスタに先読みの要否を設定することで行われる。このレジスタの設定は、オペレーティングシステム (O S) が各プロセッサ 1 1, 1 2 で実行すべきプログラムを判断して、プログラムに応じて設定する。尚、設定するレジス

タは、プロセッサ内に設けられている必要はなく、プロセッサと1:1で設けられたSB内のレジスタであれば良い。更に、先読みの要否をプロセッサ毎に設定するのではなく、全てのプロセッサに対して同じ設定とする場合には、プロセッサ毎にレジスタを設ける必要がないことは、言うまでもない。

【0026】

つまり、OSは、どのプログラムをどのSB 1-1~1-N内のどのプロセッサ 11, 12に実行させるかを決定する。そこで、OSがプログラムを実行すべきプロセッサ 11, 12を決定する際に、上記レジスタの設定を行う。

図2の上部に示すように、あるデータ群を使用するプログラムAが、あるSB内の1つのプロセッサ 11でのみ実行され、同じSB内の他のプロセッサ 12及び他のSB内の全てのプロセッサ 11, 12はスリープ状態又はプログラムAとは全く関係のないプログラムを実行している。従って、このような場合には、プログラムAで使用するデータが、同じSB内の他のプロセッサ 12や他のSB内の全てのプロセッサ 11, 12内のキャッシュメモリ 11a, 12aにある可能性は非常に低く、データの先読みが成功する確率は非常に高い。そこで、このような場合には、各SB内の各プロセッサ 11, 12に対応するレジスタに、データの先読みが必要である旨をOSにより設定しておけば良い。プロセッサと1:1でレジスタが設けられている場合、プログラムAが走っているプロセッサに対応するレジスタだけに先読みを設定すれば良い。尚、夫々のレジスタの設定に応じて、先読みを行うか否かの情報が命令に付加される。

【0027】

他方、図2の下部に示すように、1又は複数のSB内の複数のプロセッサ 11, 12で1つのプログラムBを実行する場合や、複数のプロセッサ 11, 12でデータベースのデータを共有して使用する場合等には、使用するデータが他のプロセッサ 11, 12内のキャッシュメモリ 11a, 12aにある可能性が高い。そこで、このような場合には、各SB内の各プロセッサ 11, 12に対応するレジスタに、データの先読みが不要である旨をOSにより設定しておけば良い。

【0028】

上記の如き、プロセッサ 11, 12に対応するレジスタの設定に応じて、先読

みデータの転送と通常のデータ転送とを区別して、先読みデータの優先度の方を低く設定する。例えば、データの先読みが設定されたプロセッサから発行されるリード要求に対しては、「先読みが必要」であることを示す情報を付加しておく。メモリアクセスの際には、この情報に基づいて、データ先読みの優先度を通常のメモリアクセスの優先度より下げてデータ転送を行うことができる。又、先読みデータにも同様の情報を付加しておくことで、先読みデータの優先度を通常のデータ転送の優先度より下げてデータ転送を行うことができる。従って、バス調停装置 20, 35 は、要求やデータに付加された情報に基づいて、バス 3 の使用权を決定することができる。尚、先読みの要否を示す情報は、例えば先読みフラグ等の形で要求やデータに付加することができる。

【0029】

図 3 は、本実施例の動作タイミングを説明するための図である。同図中、(M) はリード要求を行った SB 1-1、(S) は要求されたアドレスのメモリを持つ SB 1-2、(全) は全ての SB 1-1 ~ 1-N、AC は XB 2 のアドレス広報装置 32、CC は XB 2 のキャッシュ情報広報装置 33、DC は XB 2 のデータ転送装置 31 を示す。同図からもわかるように、データ先読み実行時と、データ先読みを実行しない時との差は、T となる。

【0030】

ところで、XB 2 のアドレス広報装置 32 からアドレスが広報されて各 SB 1-1 ~ 1-N に到着したタイミングでメモリリードを開始するモードと、各 SB 1-1 ~ 1-N 内のキャッシュ状態を読み出して確認した後にメモリリードを開始するモードとで、別々なレジスタ、即ち、第 1 のレジスタと第 2 のレジスタを、各プロセッサ 11, 12 に対して設けても良い。

【0031】

この場合、第 1 及び第 2 のレジスタが両方とも有効な設定となっていれば、各 SB 1-1 ~ 1-N へのアドレス到着時にデータの先読みアクセスを開始し、開始できるまでにキャッシュ情報が読み出されたならこれを確認した後に、再びメモリアクセスを開始する。アドレス到着時にデータの先読みアクセスを開始できない場合には、開始できるまでメモリアクセスを繰り返し行う。又、キャッシュ

情報の確認後に再びメモリアクセスを開始できない場合にも、開始できるまでメモリアクセスを繰り返し行う。

【 0 0 3 2 】

通常は、第 1 及び第 2 のレジスタの一方のみを有効に設定することはないが、次のように設定条件を決定しても良い。つまり、第 1 のレジスタのみが有効に設定されている場合には、データの先読みアクセス開始が成功するまでにキャッシュ情報が読み出されれば、その時点で先読みアクセスを中断する。他方、第 2 のレジスタのみが有効に設定されている場合には、キャッシュ情報を読み出して確認した後にデータの先読みアクセスを開始する。実際のマルチプロセッサシステムの試験を行って性能を調べることにより、第 1 及び第 2 のレジスタをどのように設定するのが、マルチプロセッサシステムの性能を最大限生かすことができるかを知ることができる。

【 0 0 3 3 】

図 4 は、上記の如き本実施例の動作の概略を示すフローチャートである。同図中、(M) はリード要求を行った S B 1 - 1、(S) は要求されたアドレスのメモリを持つ S B 1 - 2、X B は X B 2 を示す。

図 4 において、ステップ 1 0 0 では、S B (M) 内のプロセッサでミスヒットが発生すると、アドレス、リクエスト等を含むリード要求を発行する。又、全体制御回路 1 7 は、S B (M) 内の他のプロセッサからの要求があれば、調停を行う。これにより、全体制御回路 1 7 からは、アドレス、リクエスト、要求を発行したプロセッサの I D、先読み要否等を含むリード要求が、X B に対して出力される。

【 0 0 3 4 】

ステップ 1 0 1 では、X B がアドレス広報を行い、複数の S B からの要求があれば、調停を行う。

ステップ 1 0 2 では、各 S B において、キャッシュメモリのタグ情報を調査すると共に、リード要求の要求アドレスのメモリを持つ S B であるか否かを判定し、キャッシュ情報を X B へ送る。又、S B (S) においては、メインメモリ 1 6 をアクセスして読み出した先読みデータを、データ保持バッファ 2 1 に保持する

。データ保持バッファ 2 1 内の先読みデータは、上記 S B (M) 内のリード要求を発行したプロセッサの I D である転送先情報、先読みデータフラグ等と共に、X B に転送される。つまり、ステップ 1 0 2 は、メインメモリ 1 6 内に要求アドレスがあるか否かを見る処理と、タグ情報を見てキャッシュメモリの状態を知る処理との、2 つの処理を行う。この時に、要求アドレスのメモリを持っているものが、S B (S) である。

【 0 0 3 5 】

ステップ 1 0 3 では、X B が S B から送られてきた順にデータをデータ保持バッファ 3 4 に一旦保持し、S B (M) に対して出力する。複数の S B からデータが同時に送られてくれば、調停を行う。更に、先読みデータの転送の優先度は、通常のデータ転送の優先度より低く設定する。又、X B は、全ての S B からのキャッシュ情報を選択、若しくは、マージして、キャッシュ情報の広報を、各 S B に対して行う。

【 0 0 3 6 】

ステップ 1 0 4 では、先読みが成功すれば、X B からの先読みデータを S B (M) 内のデータ保持バッファ 2 1 に保持し、リード要求を発行したプロセッサによる取り込みを可能とする。

従って、本実施例におけるリード要求に基づくデータの先読みは、次の手順で行われる。

【 0 0 3 7 】

ステップ S 1 : 上記マルチプロセッサシステムにおいて、要求アドレスのメモリを持つ S B は、全キャッシュメモリの状態を示すキャッシュ情報が得られる前にリード、即ち、データの先読みを開始する。

ステップ S 2 : ステップ S 1 を行う際、先読みデータは、転送パスの途中で保持する。転送パスの途中とは、後述する如く、リード要求の要求アドレスのメモリを持つ S B 内のデータ保持バッファ 2 1、X B 内のデータ保持バッファ 3 4、又は、リード要求を発行した S B 内のデータ保持バッファ 2 1 である。

【 0 0 3 8 】

ステップ S 3 : X B において集めた各 S B からのキャッシュ情報に基づき、

転送パスの途中で保持された先読みデータが有効であることを確認した後、データ転送を続行する。

ステップ S 4 : 要求アドレスを持つ S B は、この S B 内のキャッシュメモリの内容に関わらずデータの先読みを開始するか、或いは、キャッシュメモリの内容を確認した後にデータの先読みを開始する。

【 0 0 3 9 】

ステップ S 5 : データの先読みが開始できない場合には、上記ステップ S 3 又はステップ S 4 を少なくとも 1 回繰り返す。尚、ステップ S 3 及びステップ S 4 は同時に行われないように、フラグを設定しておく。又、上記の如く、データの先読みの要否は、実行するプログラムに応じて、プロセッサに対応するレジスタに設定しておく。

【 0 0 4 0 】

このように、大規模なマルチプロセッサシステムでは、リード要求を発行したプロセッサのできるだけ近くに、このプロセッサが必要とするデータを転送しておくことと、より早いタイミングでデータの先読みを開始することが重要である。そこで、本実施例では、データの転送パスの途中にバッファを設け、通常のデータ転送の妨げとならないように、先読みデータの転送の優先度は通常のデータ転送の優先度より下げてバスの調停を行う。又、先読みデータは、データの転送パス上を順次転送して、徐々にリード要求を発行したプロセッサに近づけて行く。

【 0 0 4 1 】

他方、データの先読み開始のタイミングは、全ての S B にアドレスが広報され、このアドレスを受け取った直後のタイミングに設定する。つまり、データの先読み開始のタイミングは、S B 内のキャッシュメモリの状態を示すタグ情報を調べるタイミングと同時とする。これにより、非常に早くデータの先読みを開始できる。又、他からのメモリアクセスの混み具合に応じて、前記タイミングでデータの先読みが開始できない場合には、データの先読み開始を繰り返す。更に、データの先読みが、メモリに受け入れられる前にキャッシュメモリのタグ情報を調べ終わった場合には、タグ情報から S B 内のキャッシュメモリにデータがあるか

否かの判断を行い、データが無い場合にはデータの先読み開始を繰り返す。

【0042】

次に、本発明になるマルチプロセッサシステムの第2実施例を説明する。図5は、マルチプロセッサシステムの第2実施例を示すブロック図である。マルチプロセッサシステムの第2実施例では、本発明になるメモリアクセス方法の第2実施例を採用する。

図5において、マルチプロセッサシステムは、複数のノード50-1～50-4からなる。各ノード50-1～50-4は、複数のSB1-1、1-2と、XB2と、これらを接続するバス3とからなる。SB1-1、1-2、XB2及びバス3の構成は、上記第1実施例と同じである。隣接するノードのXB2は、バス4により接続されている。

【0043】

尚、ノードの数は4に限定されるものではなく、又、各ノード内のXBの数も2に限定されるものではない。

本実施例では、1つのノード内においてデータの先読みが行われる場合と、複数のノードにまたがってデータの先読みが行われる場合とがある。前者の場合の動作は、上記第1実施例と同じである。後者の場合の動作は、次の通りである。

【0044】

つまり、例えばノード50-1内のSB1-1からリード要求があり、要求アドレスのメモリが隣接するノード50-2内のSB1-2にある場合には、先読みデータがノード50-2内のSB1-2から順次ノード50-2内のXB2及びバス4、要求元のノード50-1内のXB2を介してノード50-1内のSB1-1に転送される。従って、本実施例においても、リード要求を行ったプロセッサにできるだけ近い場所で先読みされたデータを保持することができる。

【0045】

以上、本発明を実施例により説明したが、本発明は上記実施例に限定されるものではなく、本発明の範囲内で種々の変形及び改良が可能であることは、言うまでもない。

【0046】

【発明の効果】

本発明によれば、リード要求を行ったプロセッサにできるだけ近い場所に先読みされたデータを保持して、通常のデータ転送を妨害することなく、先読みの本来のメリットを生かすことができ、マルチプロセッサシステム全体としての性能を向上可能となる。

【図面の簡単な説明】

【図 1】

本発明になるマルチプロセッサシステムの第 1 実施例を示すブロック図である。

【図 2】

データの先読みを行うか否かの設定を説明するための図である。

【図 3】

第 1 実施例の動作タイミングを説明するための図である。

【図 4】

第 1 実施例の動作の概略を示すフローチャートである。

【図 5】

本発明になるマルチプロセッサシステムの第 2 実施例を示すブロック図である。

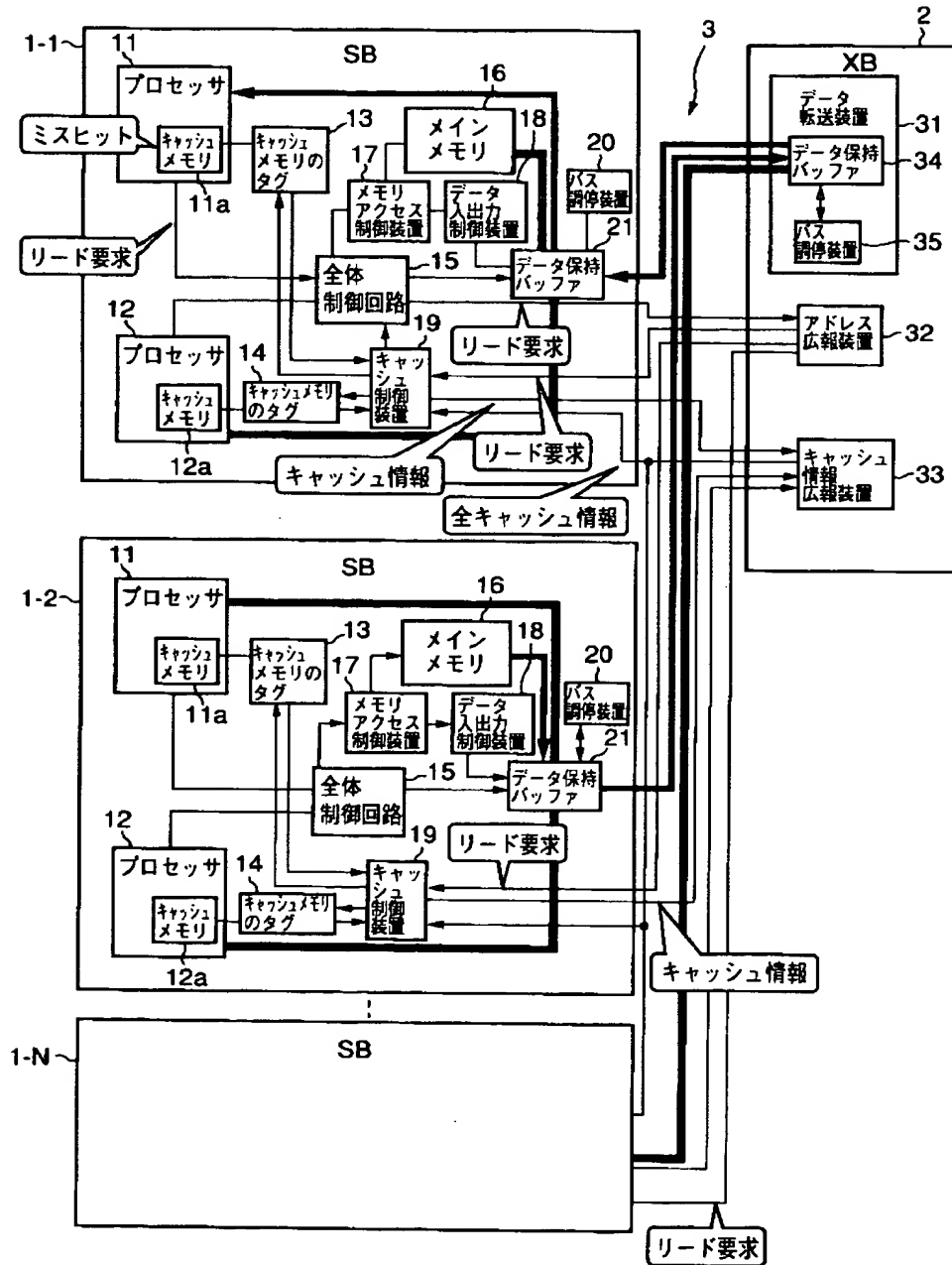
【符号の説明】

- 1 - 1 ~ 1 - N システムボード (SB)
- 2 クロスバボード (XB)
- 3, 4 バス
- 1 1, 1 2 プロセッサ
- 1 6 メインメモリ
- 2 1, 3 4 データ保持バッファ
- 5 0 - 1 ~ 5 0 - 4 ノード

【書類名】 図面

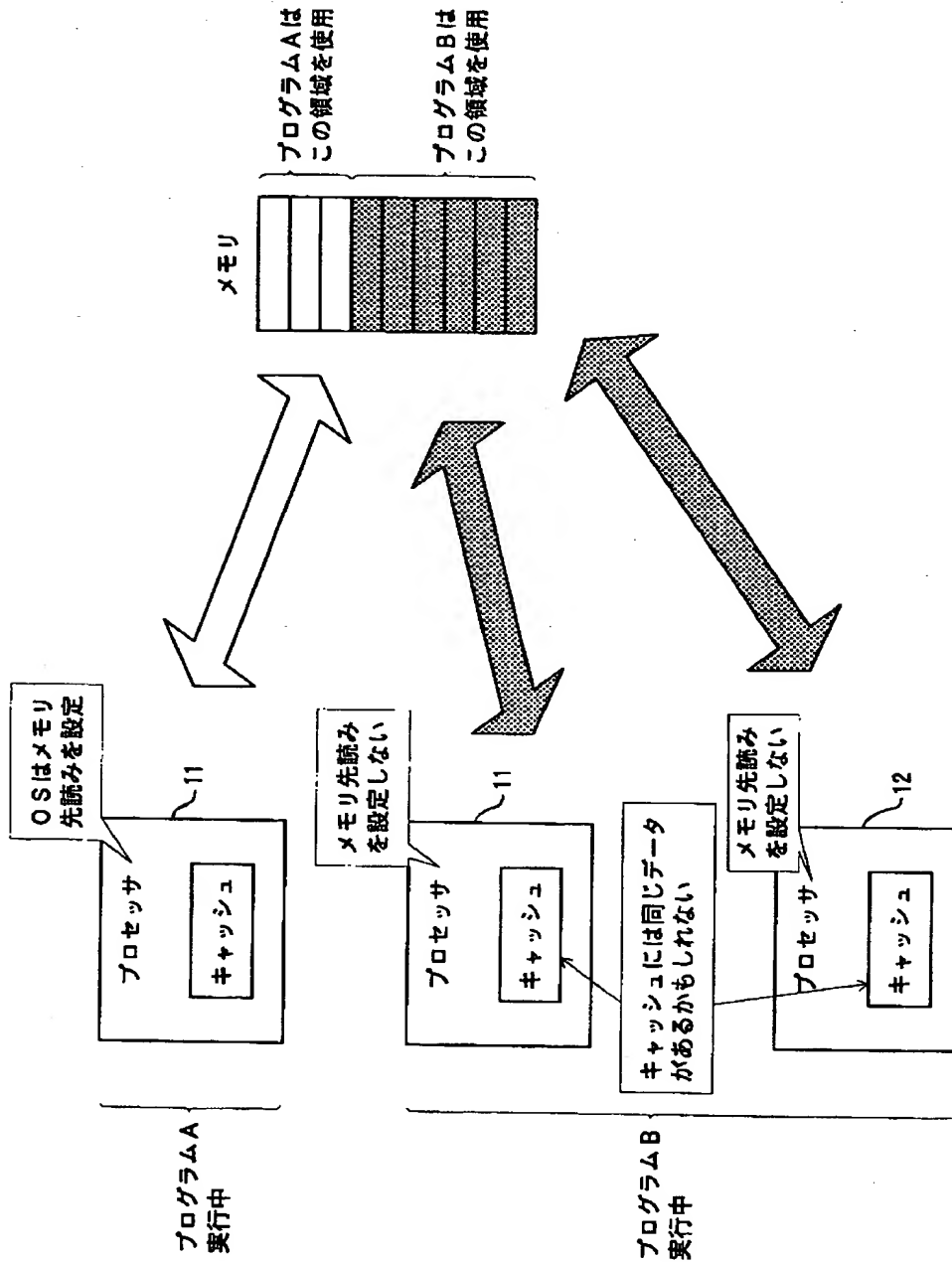
【図 1】

本発明になるマルチプロセッサシステムの第1実施例を示すブロック図



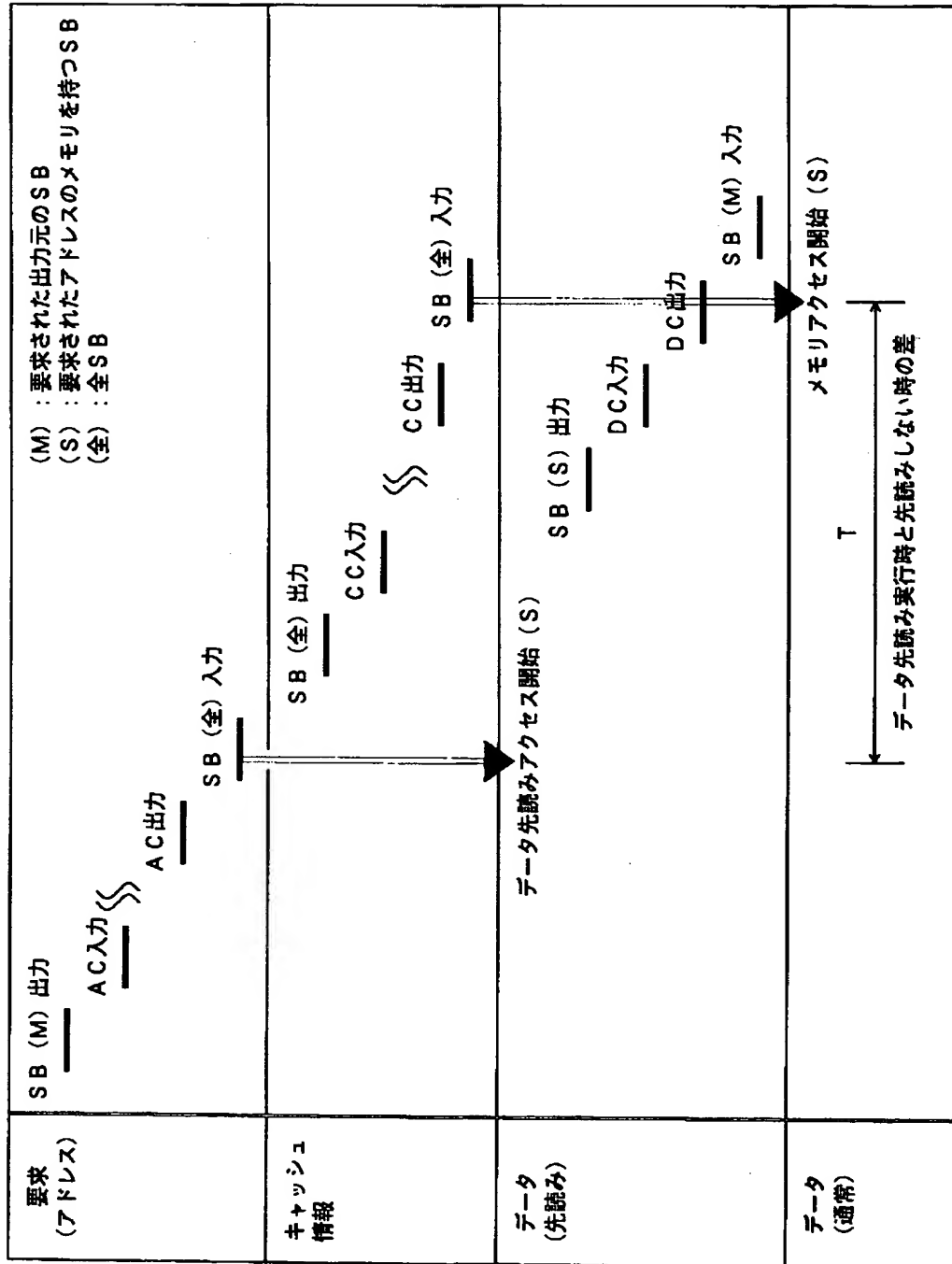
【図 2】

データの先読みを行うか否かの設定を説明するための図



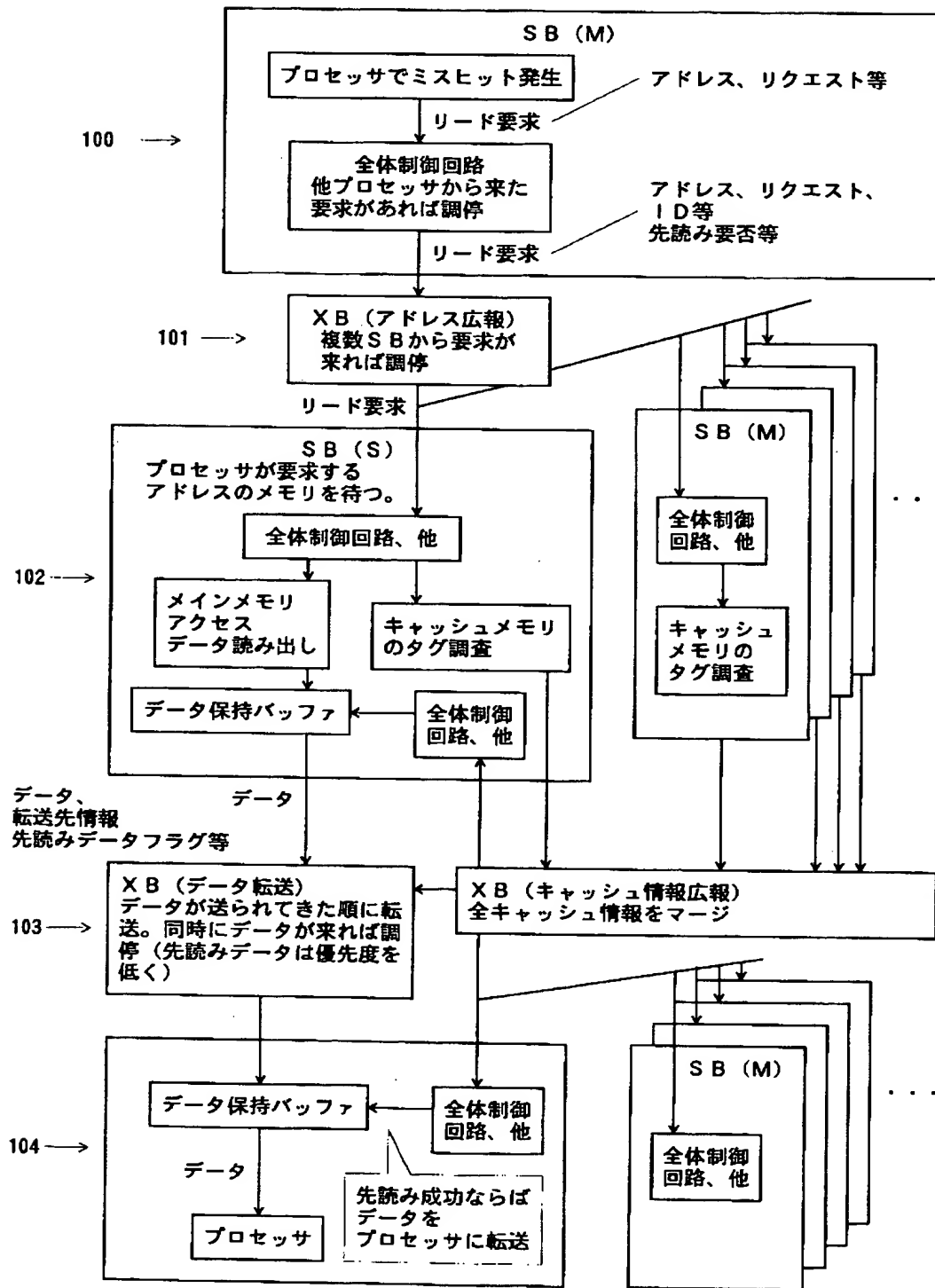
【図 3】

第 1 実施例の動作タイミングを説明するための図



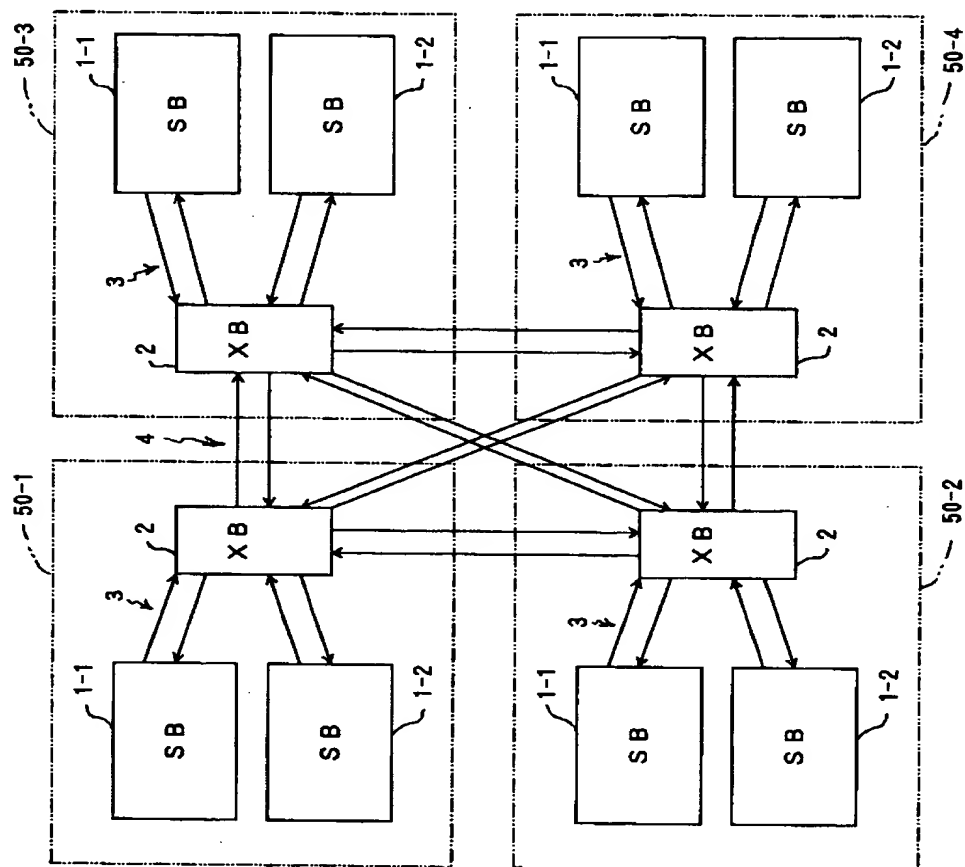
【図 4】

第 1 実施の動作の概略を示すフローチャート



【図 5】

本発明になるマルチプロセッサシステムの第 2 実施例を示すブロック図



【書類名】 要約書

【要約】

【課題】 本発明はメモリアクセス方法及びマルチプロセッサシステムに関し、リード要求を行ったプロセッサにできるだけ近い場所に先読みされたデータを保持して、通常のデータ転送を妨害することなく、先読みの本来のメリットを生かすことができ、マルチプロセッサシステム全体としての性能を向上可能とすることを目的とする。

【解決手段】 データを保持するバッファと、各々がデータを一時的に保持するキャッシュメモリを有する複数のプロセッサを搭載したシステムモジュールが複数台、クロスバモジュールを介して接続された構成のマルチプロセッサシステムにおけるメモリアクセス方法において、任意の 1 台のシステムモジュール内のプロセッサからのリード要求に応答して、この任意の 1 台のシステムモジュール以外のシステムモジュールから先読みされたデータを、前記クロスバモジュール内のバッファに保持するステップを含むように構成する。

【選択図】 図 1

出 願 人 履 歴 情 報

識別番号 [000005223]

1. 変更年月日	1996年 3月26日
[変更理由]	住所変更
住 所	神奈川県川崎市中原区上小田中4丁目1番1号
氏 名	富士通株式会社